JOURNAL OF BACTERIOLOGY, Feb. 1993, p. 716–722
0021-9193/93/030716-07$02.00/0

Vol. 175, No. 3

Effects of Consecutive AGG Codons on Translation in *Escherichia coli*, Demonstrated with a Versatile Codon Test System

ALAN H. ROSENBERG,[1]* EMANUEL GOLDMAN,[2] JOHN J. DUNN,[1]
F. WILLIAM STUDIER,[1] AND GEOFFREY ZUBAY[3]

Biology Department, Brookhaven National Laboratory, Upton, New York 11973[1]; Department of Microbiology and Molecular Genetics, University of Medicine and Dentistry of New Jersey–New Jersey Medical School, Newark, New Jersey 07103[2]; and Fairchild Center for Biological Sciences, Columbia University, New York, New York 10027[3]

Received 27 July 1992/Accepted 23 November 1992

A system for testing the effects of specific codons on gene expression is described. Tandem test and control genes are contained in a transcription unit for bacteriophage T7 RNA polymerase in a multicopy plasmid, and nearly identical test and control mRNAs are generated from the primary transcript by RNase III cleavages. Their coding sequences, derived from T7 gene *9*, are translated efficiently and have few low-usage codons of *Escherichia coli*. The upstream test gene contains a site for insertion of test codons, and the downstream control gene has a 45-codon deletion that allows test and control mRNAs and proteins to be separated by gel electrophoresis. Codons can be inserted among identical flanking codons after codon 13, 223, or 307 in codon test vectors pCT1, pCT2, and pCT3, respectively, the third site being six codons from the termination codon. The insertion of two to five consecutive AGG (low-usage) arginine codons selectively reduced the production of full-length test protein to extents that depended on the number of AGG codons, the site of insertion, and the amount of test mRNA. Production of aberrant proteins was also stimulated at high levels of mRNA. The effects occurred primarily at the translational level and were not produced by CGU (high-usage) arginine codons. Our results are consistent with the idea that sufficiently high levels of the AGG mRNA can cause essentially all of the $tRNA^{AGG}$ in the cell to become sequestered in translating peptidyl-$tRNA^{AGG}$-mRNA-ribosome complexes stalled at the first of two consecutive AGG codons and that the approach of an upstream translating ribosome stimulates a stalled ribosome to frameshift, hop, or terminate translation.

Most amino acids are encoded by more than one codon, and frequencies of use of synonymous codons tend to reflect the relative abundance of the tRNAs that recognize them (7, 29). Synonymous codons may be translated at different rates (16, 17), but what role, if any, codon usage may play in controlling the expression of individual proteins in normal cells is not clear (reviewed by Andersson and Kurland [1]).

Several examples suggest that the arginine codons AGG and AGA can have pronounced effects on the translation of cloned genes in *Escherichia coli*. These two codons are the least frequently used in *E. coli* (29), and the tRNAs that recognize them are among the least abundant (7, 9). Four or five consecutive AGG codons can substantially reduce the expression of test proteins (5, 12), although six AGG codons separated from each other by single GCG (alanine) codons did not (26). Even one AGG codon inserted in the leader peptide of a bacterial attenuator apparently retarded the movement of ribosomes sufficiently to increase the efficiency of attenuation (3). Two consecutive AGG or AGA codons stimulated a high level of frameshifting, which was suppressed when additional tRNA capable of recognizing these codons was provided (18, 19). Some eukaryotic genes that use high levels of AGA and AGG codons (3.0 to 4.9%) were expressed poorly in *E. coli* itself but were expressed well when additional tRNA was supplied (4). On the other hand, a 153-codon fragment of the influenza virus cap-binding protein PB2 that contains four AGG and seven AGA codons (7.2%) was expressed very efficiently without supplementary tRNAs (14).

A thorough understanding of how specific codons can affect translation should improve our understanding of the translation apparatus of *E. coli* and might also help to solve problems encountered in expressing cloned genes. To be able to study the effects of specific codons systematically, we developed a new codon test system with many desirable features. We describe this system and use it to analyze the effects of inserting two to five consecutive AGG codons at three different positions in a test gene.

MATERIALS AND METHODS

Codon test plasmids. Plasmids pCT1, pCT2, and pCT3 (Fig. 1) were assembled from elements of pET vectors (13, 24) and of T7 DNA (6). Elements were assembled by using a natural *Xba*I site (T'CTAGA, where the primer indicates the position of cleavage) between the ϕ*10* promoter and the s*10* translation initiation region for the gene *10* major capsid protein of T7, a natural *Nde*I site (CA'TATG) at the s*10* initiation codon, and newly introduced *Xba*I or *Nhe*I (G'CTAGC) sites at the ends of elements. Plasmids were assembled in modular fashion by fusions of compatible ends of DNA fragments from plasmids carrying individual elements or combinations of elements. Most of the fusions were between fragments that had one end at the *Pst*I site in the

* Corresponding author.

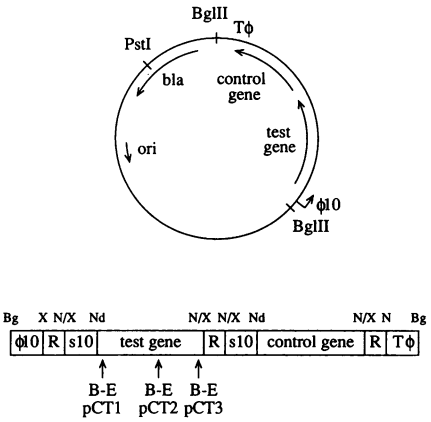

FIG. 1. Arrangement of elements in the codon test vectors pCT1, pCT2, and pCT3. φ*10* is a promoter for T7 RNA polymerase, R is an RNase III cleavage site, s*10* is a strong translational start, and Tφ is a transcriptional terminator for T7 RNA polymerase. The test and control genes are derived from T7 gene *9*, as described in the text. A *Bam*HI-*Eco*RI site allows insertion of codons after codon 13, 223, or 307 in the test genes of pCT1, pCT2, and pCT3, respectively. The control gene has a 45-codon deletion that allows its mRNA and protein to be distinguished from the test mRNAs and proteins by gel electrophoresis. Restriction sites: B-E, *Bam*HI-*Eco*RI cloning site; Bg, *Bgl*II; Nd, *Nde*I; N, *Nhe*I; X, *Xba*I; N/X, *Nhe*I-*Xba*I fusion.

β-lactamase gene and the other at an *Xba*I, *Nhe*I, or *Nde*I site at an end of one of the desired elements.

Individual elements were generated by restriction cleavage and addition of linkers or by amplification by polymerase chain reaction (8) with primers used to introduce appropriate linking sequences. The positions of the elements in the sequence of T7 DNA were: φ*10* promoter, bp 22880 to 22928 (bp −23 to +26 relative to the RNA start); R*0.3* RNase III cleavage site, bp 823 to 906 (RNase III cuts once, between nucleotides 890 and 891); s*10* translation initiation region, bp 22925 to 22968; coding sequence of T7 gene *9* (307 codons), bp 21949 to 22870 (with the termination codon changed to TAA from TGA); and Tφ transcription terminator, bp 24130 to 24228 (transcription terminates at bp 24209). The transcription unit in the plasmids is flanked by *Bgl*II sites (A'GATCT) (Fig. 1). The sequence upstream of φ*10* was from pET-3b, and the sequence downstream of Tφ was from pET-3 (24). The region (386 bp) between the downstream *Bgl*II site and the *Eco*RI site (G'AATTC) present in pET-3 was deleted, eliminating the *Eco*RI site and leaving the junction sequence AGATCTAATTC. The *Bgl*II sites flanking the transcription unit allow it to be isolated as a single DNA fragment for transfer to other vectors or the chromosome.

To make the control gene, 45 codons (codons 151 to 195) of the cloned gene *9* coding sequence were deleted by removing the DNA between *Rsa*I sites (bp 22400 and 22535 of T7 DNA). To make the test genes, unique *Nsi*I, *Alw*NI, and *Cla*I sites in the gene *9* coding sequence were modified to create *Bam*HI and *Eco*RI sites for directional insertion of identical sets of test codons at three different positions. The sequence CTGGATCCGAAGCTTGAATTC was inserted immediately following codon 10, 220, or 304 in the 307-codon gene *9* coding sequence. The inserted sequence specifies Leu-Asp-Pro-Lys-Leu-Glu-Phe, with no low-usage codons. Codons 9 and 10 (Tyr-Ala), codon 220 (Ser), or codon 304 (Asp) were repeated after the inserted sequence. The *Bam*HI and *Eco*RI sites are separated by 7 bp, so that cutting with both enzymes will easily go to completion. Codons to be tested replace the Lys-Leu codons. The correctness of each replacement that we made was verified by nucleotide sequencing.

The proper functioning of each element of the transcription unit was tested in vivo in the precursor and final plasmids, and the predicted nucleotide sequences across junctions between the different elements were verified. Plasmid pCT1 is predicted to be 6,258 bp long, and pCT2 and pCT3 are each predicted to be 6,255 bp long. The predicted nucleotide sequences of these plasmids are available from A. H. Rosenberg.

Expression and analysis of RNAs and proteins. T7 RNA polymerase was supplied from a chromosomal copy of the polymerase gene under control of the isopropylthiogalactopyranoside (IPTG)-inducible *lac*UV5 promoter in HMS174 (DE3) (F^- *recA* r_{K12}^- m_{K12}^- Rif^r) or BL27(DE3), a *recA* derivative of BL21(DE3) (F^- *ompT* r_B^- m_B^-) (23). BL27 (DE3) was made by Paresh Shrimankar (Case Western University) by transduction of BL21(DE3) to Tet^r with P1 *vir* phage grown on donor strain RM1981 (*recA56 srl-300*::Tn*10*). *recA* host cells were used to reduce the possibility of recombination between the tandem genes in the test plasmids. Expression from each codon test plasmid (Amp^r) was done in the presence of the compatible plasmid pLysS (Cm^r), which reduces basal transcription by T7 RNA polymerase but allows high-level expression after induction (21).

Growth of cultures, induction of expression, labeling of proteins with [^{35}S]methionine and RNAs by $^{32}PO_4$, and analysis by gel electrophoresis in the presence of sodium dodecyl sulfate on 10 to 20% gradient polyacrylamide gels were done essentially as described before (23, 24). Northern (RNA blot) analysis of mRNAs was done by the procedures described by Sambrook et al. (15).

RESULTS

Codon test system. The system for measuring the possible effects of specific codons on translation contains tandem test and control genes expressed from a single transcription unit for T7 RNA polymerase (Fig. 1). Codons to be tested are inserted into the upstream gene, and their effects on translation are detected as changes in production of the test protein relative to the control protein.

The primary transcript, from the φ*10* promoter to the Tφ terminator for T7 RNA polymerase, contains three identical RNase III cleavage sites, which separate the test and control mRNAs from each other and from the flanking transcription signals. The two mRNAs produced by cleavage at these sites are identical except for the site for inserting codons into the test gene and a 45-codon deletion in the control gene, to allow the two mRNAs (and proteins) to be separated by gel electrophoresis. Independent translation of test and control mRNAs avoids possible interactions between two coding sequences in the same mRNA, and RNase III cleavage leaves a strong stem-loop structure at the 3' end, which should stabilize the mRNAs (6, 11).

The test and control coding sequences are derived from gene *9* of T7, which specifies a capsid assembly protein of 306 amino acids (after removal of the initial methionine). Each coding sequence is fused to the efficient upstream translation initiation region for the T7 major capsid protein, referred to as s*10* (13). The gene *9* protein is expressed well, highly soluble, and not particularly toxic to the host cell (23).

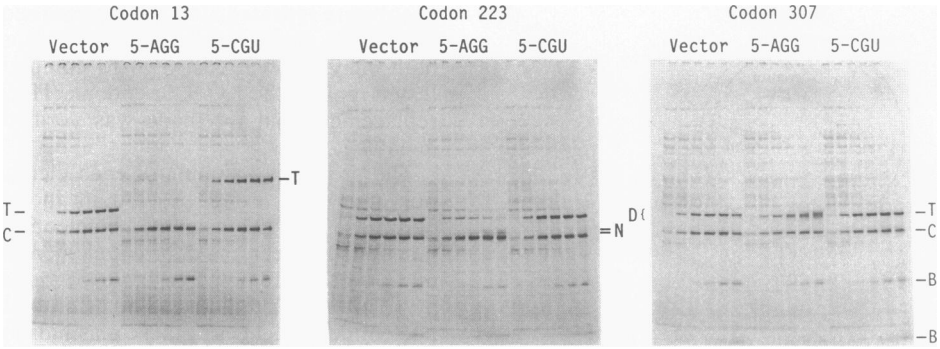

FIG. 2. Effects of five consecutive low-usage (5-AGG) or high-usage (5-CGU) arginine codons on synthesis of test proteins. Cultures of HMS174(DE3)/pLysS carrying different test plasmids were grown in M9 medium plus ampicillin (20 μg/ml) and chloramphenicol (25 μg/ml) at 37°C, labeled for 2 min with [^{35}S]methionine immediately before and 5, 10, 20, 30, and 60 min after induction of T7 RNA polymerase by 0.4 mM IPTG, and analyzed for labeled proteins by gel electrophoresis and autoradiography. The equivalent of 20 μl of culture was loaded in each gel track. In each panel, sets of samples are shown for the vector and for the vector with insertions of either five consecutive low-usage (AGG) codons or five consecutive high-usage (CGU) codons, as indicated. In the left panel, plasmids were pCT1, pAR5021, and pAR5023, and codon insertions were after codon 13; in the middle panel, plasmids were pCT2, pAR5057, and pAR5059, and insertions were after codon 223; in the right panel, plasmids were pCT3, pAR5025, and pAR5027, and insertions were after codon 307. Positions are indicated for the test (T) and control (C) proteins, smaller bands related to the control protein (B), two new bands produced when AGG codons were inserted after codon 223 (N), and a diffuse band produced when AGG codons were inserted after codon 307 (D).

Its codon composition is typical of genes that are highly expressed in *E. coli* and includes no AGG or AGA codons. Test and control proteins are easily detected by gel electrophoresis of whole-cell extracts followed by staining or autoradiography (each protein contains six internal methionines). The analysis does not rely on enzymatic activity, which could be affected by the amino acids inserted into the test protein.

Three separate test vectors were made, allowing codons to be inserted immediately following codon 13, 223, or 307 of the test gene in plasmids pCT1, pCT2, and pCT3, respectively. The insertion site after codon 307 is six codons from the termination codon. In each vector, codons are inserted in the same local context.

T7 RNA polymerase is supplied from an IPTG-inducible gene in the chromosome of the host cell (23). Promoters recognized by T7 RNA polymerase do not occur in *E. coli* DNA, so that all transcription by T7 RNA polymerase will be directed toward the codon test plasmid. T7 RNA polymerase is very processive, and equal amounts of test and control mRNA should be produced even if translation of the test mRNA is completely blocked. A wide range of mRNA levels is possible: the activity of T7 RNA polymerase can be modulated by T7 lysozyme, a natural inhibitor (21), and competing mRNAs can be eliminated by selectively inhibiting the host RNA polymerase with rifampin.

This system should facilitate systematic investigation of the effects of specific codons on translation. Producing test and control mRNAs and proteins in the same cell controls for many sources of experimental variation. The availability of three cloning sites permits analysis of position-dependent effects. Comparing synonymous codons should distinguish effects of codon usage from effects on the protein, and measuring levels of test and control mRNA can test whether effects occur at the transcriptional level.

Effects of codon insertions on protein synthesis. We used the codon test system to analyze effects of the AGG arginine codon, which has been shown to reduce protein synthesis or cause frameshifting in at least some situations (5, 12, 19, 26). Five consecutive AGG (low-usage) or CGU (high-usage) arginine codons were inserted into each of the three test genes, and protein synthesis was measured as a function of time after induction of T7 RNA polymerase (Fig. 2).

Nearly equal amounts of test and control proteins were produced from the codon test vectors pCT1, pCT2, and pCT3 themselves and when five consecutive CGU codons had been inserted (Fig. 2). Two unexpected effects were observed, neither of which has much consequence for the use of these vectors to measure the effects of codon usage.

(i) Small amounts of two more rapidly migrating bands were apparent in pulse-labeled samples, particularly at later times after induction (bands B in Fig. 2). Induction of test and control genes in separate plasmids showed that these bands arise primarily from the control gene, but we have not determined what part of the coding sequence they represent or how they are generated.

(ii) The mobility of the test protein decreased substantially when arginine codons were inserted after codon 13 (Fig. 2; also see Fig. 5). Insertion of some amino acids besides arginine at this site also decreased protein mobility, but insertion of others did not (not shown). None of the insertions after codon 223 or 307 had such an effect. Perhaps these unexpected mobility changes are somehow related to known changes in the relative mobility of gene *9* protein with changes in gel concentration (22).

As expected, insertion of five consecutive AGG (low-usage) codons into the test gene had little effect on expression of the control gene but significantly affected expression of the test gene. The effects depended on where the codons had been inserted (Fig. 2).

Insertion of the AGG codons after codon 13 severely inhibited synthesis of the test protein, which was produced at a detectable rate for less than 20 min after induction. Possible products of termination or frameshifting near codon 13 would not have been detected.

Insertion of the AGG codons after codon 223 also strongly inhibited production of full-length test protein but not as severely as insertion after codon 13. The rate decreased at later times, but in contrast to insertion after codon 13, synthesis of full-length protein continued for at least 60 min after induction. Significant amounts of at least two smaller bands that migrated just behind the control protein also

began to appear by 20 min after induction (N bands in Fig. 2). These new proteins are likely products of frameshifting (19) or possibly other termination events by ribosomes stalled at the low-usage codons. Shifting into the other reading frames at the string of AGG codons would be expected to produce proteins of 232 and 235 amino acids. The two new proteins migrate as though they are longer than the control protein of 261 amino acids (which has an internal deletion), but their mobility is similar to that of gene *9* protein truncated near codon 223 (not shown).

Insertion of the AGG codons after codon 307, only six codons from the end of the protein, also inhibited synthesis of full-length test protein. The extent of inhibition was difficult to determine because the band became obscured by increasing amounts of a diffuse band of protein that began to appear by 20 min after induction (marked D in Fig. 2). Full-length test protein should contain 317 amino acids; frameshifting at the string of AGG codons would be expected to generate proteins of 319 and 323 amino acids, but the diffuseness of the band suggests additional events. By 30 min after induction, the rate of synthesis of proteins in the diffuse band was about equivalent to or slightly greater than that of control protein, as confirmed by measuring the amount of label in bands cut from the gel (not shown).

The combined production of full-length test protein and novel proteins seemed to be greater when the AGG codons were inserted after codon 307 than after codon 223 (Fig. 2; also see Fig. 5). This apparent difference may be deceptive, however, since products of termination or frameshifting near codon 223 would incorporate either four or five rather than six labeled methionines, and some of these products might have been obscured under the band of control protein.

Effects of codon insertions on RNA synthesis. The codon test vectors were designed so that the primary transcripts would be processed by RNase III to produce discrete test and control mRNAs that can be resolved by gel electrophoresis. Samples of cultures similar to those analyzed in Fig. 2 were analyzed for accumulations of test and control mRNAs at different times after induction, by probing Northern blots with a nick-translated DNA probe prepared from the exact coding sequence of T7 gene *9*.

Neither test nor control mRNAs were detected before induction, but both were detected 10, 30, and 60 min after induction (Fig. 3). For each cloning site, there was at least as much AGG test mRNA as control mRNA or CGU test mRNA. Therefore, the reductions in full-length test protein caused by AGG codons (Fig. 2) are almost certainly due to effects at the level of translation.

Minor additional bands were evident in the original autoradiograms, some of which may not be easy to see in the reproduction in Fig. 3. In samples taken 30 and particularly 60 min after induction, a small fraction of the label was seen at positions expected for RNAs that had not been cut at one or more of the RNase III cleavage sites in the primary transcript (band I in Fig. 3). Apparently, more RNA was produced than could be processed.

Bands also appear in 30- and 60-min samples at positions expected for RNAs that end near the AGG codons inserted after codon 223 or codon 307 (bands F in Fig. 3). No such bands appear in the CGU samples. Perhaps a ribosome stalled at the low-usage AGG codons or a series of ribosomes backed up from this position protects the RNA from the normal degradative processes in the cell. However they are generated, these RNAs could also be sources of the new proteins (bands N and D) observed in Fig. 2. If a similar RNA was produced in which the AGG codons were inserted after codon 13, it was probably too small to be detected in this analysis.

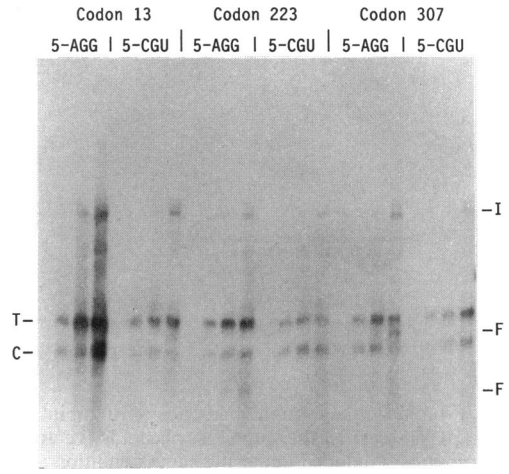

FIG. 3. Accumulations of test and control mRNAs. The same host, test plasmids (vectors omitted), and growth conditions described in the legend to Fig. 2 were used to collect parallel samples immediately before and 10, 30, and 60 min after induction with IPTG. The protein composition of one set of samples was analyzed by gel electrophoresis and staining (not shown), which confirmed the effects of codon insertions shown in Fig. 2. The RNAs of the second set of samples were subjected to gel electrophoresis on a 1.5% agarose–formaldehyde gel and then Northern blot analysis; the equivalent of 3.3 μl of culture was loaded in each gel track. An autoradiogram of a blot probed with a labeled DNA fragment specific for the test and control mRNAs is shown. The sets of samples are in the same order as in Fig. 2; test genes had five consecutive low-usage (AGG) or high-usage (CGU) arginine codons inserted after codon 13, 223, or 307, as indicated. The mobilities of the test (T) and control (C) mRNAs were consistent with their predicted lengths of 1,072 and 912 nucleotides, respectively. Positions are indicated for weak bands of incompletely processed RNA (I) and what may be RNA fragments (F) protected by ribosomes ahead of the AGG codons at codon 223 or 307.

Limited induction of mRNA. Normal induction ultimately produces large amounts of test mRNAs. Proteins produced from lower levels of test and control mRNA can be revealed by labeling in the presence of rifampin, which inhibits the host RNA polymerase but not T7 RNA polymerase. After a short period to allow host mRNAs to decay, only those mRNAs produced by T7 RNA polymerase will be available for translation. Competition from host mRNAs for limiting components of the translation apparatus, such as rare tRNAs, will also be reduced or eliminated under these conditions.

The rates of synthesis of test and control proteins in the presence of rifampin were compared for uninduced cells and after induction of T7 RNA polymerase for 5 min before the addition of rifampin (Fig. 4). The patterns of protein synthesis in the induced samples labeled 10, 25, or 55 min after rifampin addition were similar, although continuous $^{32}PO_4$ labeling of the cultures indicated that the levels of test and control mRNAs continued to increase during this period (not shown). Apparently, something other than mRNA became limiting for translation.

At the uninduced level of mRNA, test and control proteins were synthesized at comparable rates whether the test mRNA contained five consecutive AGG or CGU arginine codons at any of the three positions (Fig. 4). However, at the

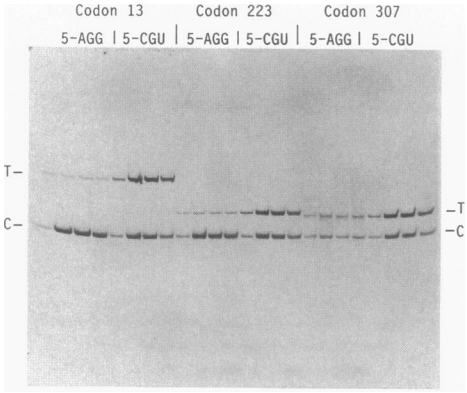

FIG. 4. Protein synthesis in the presence of rifampin. Cultures of BL27(DE3)/pLysS carried the same test plasmids (vectors omitted) and were grown and induced under the same condition as described in the legend to Fig. 2 except that a low-phosphate growth medium was used to permit $^{32}PO_4$ labeling of RNAs. Rifampin (100 μg/ml) was added to cultures before induction or after 5 min of induction with IPTG. A sample of uninduced culture was labeled for 3 min with [^{35}S]methionine 10 min after rifampin was added; samples of induced culture were labeled for 3 min at 10, 25, and 55 min after rifampin was added (that is, 15, 30, and 60 min after IPTG was added). The samples were analyzed by gel electrophoresis and autoradiography, with the uninduced culture in the leftmost track and the samples from the three time points from the induced culture in the three right-hand tracks for each set. The equivalent of 20 μl of culture was loaded in each gel track. Test genes had five consecutive low-usage (AGG) or high-usage (CGU) arginine codons inserted after codon 13, 223, or 307, as indicated.

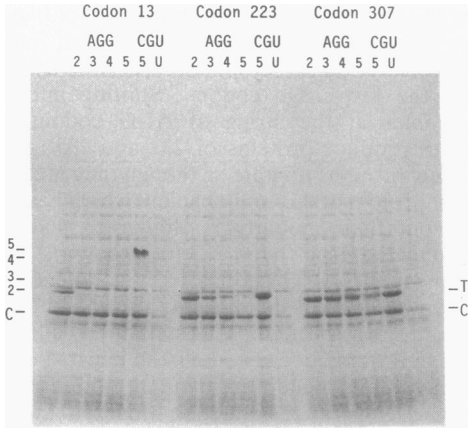

FIG. 5. Effects of two, three, four, or five consecutive low-usage (AGG) arginine codons on accumulation of test proteins. Cultures of BL27(DE3)/pLysS carrying different test plasmids were grown and induced in M9TB medium plus antibiotics at 37°C; samples were collected immediately before or 2 h after induction, and proteins were analyzed by gel electrophoresis followed by staining with Coomassie blue. The equivalent of 20 μl of culture was loaded in each gel track. Three sets of samples are shown, with insertions after codon 13, 223, or 307. The six samples in each set show, from left to right, induced cultures in which the test gene contained two, three, four, or five consecutive AGG codons and induced and uninduced (U) cultures in which the test gene contained five consecutive CGU codons. Positions of test (T) and control (C) proteins resulting from insertions after codon 223 or 307 are indicated to the right of the patterns; positions of control protein (C) and test proteins resulting from insertion of different numbers of codons after codon 13 are indicated to the left of the patterns. Test plasmids (left to right) derived from pCT1 were pAR5091, pAR5092, pAR5093, pAR5021, and pAR5023; those derived from pCT2 were pAR5094, pAR5095, pAR5096, pAR5057, and pAR5059; and those derived from pCT3 were pAR5097, pAR5098, pAR5099, pAR5025, and pAR5027.

increased levels of mRNA produced by induction, the rate of translation of test mRNAs that contained AGG codons hardly changed, whereas translation of test mRNAs that contained CGU codons increased substantially, as did translation of all but one of the control mRNAs. Small amounts of novel proteins similar to those seen in the absence of rifampin (Fig. 2) were apparent in the induced but not the uninduced samples in longer exposures of the autoradiogram shown in Fig. 4.

Insertion of different numbers of low-usage arginine codons. The pronounced effects of inserting five consecutive AGG codons led us to examine the effects of inserting fewer codons. Accumulations of test and control proteins 2 h after induction, as determined by gel electrophoresis and staining of total cell proteins, are shown in Fig. 5 for insertions of two, three, four, or five consecutive AGG codons and five consecutive CGU codons. Levels of accumulation with five consecutive codons inserted (Fig. 5) are consistent with the rates of synthesis determined by pulse labeling (Fig. 2).

The strongest effects were seen when the AGG codons were inserted after codon 13: a single AGG at this position had no detectable effect (not shown), but the accumulation of test protein decreased substantially with only two consecutive codons, decreased further with three or four codons, and was not apparent with five codons (although synthesis was detected by pulse labeling for a short time after induction; Fig. 2). The effects were less severe when insertion was made after codon 223, becoming clearly apparent only with insertion of three or more AGG codons. Insertion after codon 307 seemed to have the least effect, becoming significant only with insertion of four or five codons. The insertion of five consecutive AGG codons after codon 223 or 307 appears to have reduced the accumulation of control protein as well.

DISCUSSION

Our results and those of others (5, 12, 19, 26) show that consecutive AGG codons can modulate protein synthesis at the level of translation. We observed larger effects and a greater diversity of effects with strings of two to five AGG codons than did Chen and Inouye with β-galactosidase (5) or Spanjaard et al. with an MS2 coat protein system (18, 19), perhaps because our system generated higher levels of test mRNA, initiated protein synthesis more efficiently, or both.

The altered translation resulting from insertion of codons into an mRNA could be due to any of several possible mechanisms. Our experiments eliminate instability of the mRNA or protein as significant factors in the translational changes that we observed. Changes in secondary and tertiary structure or in specific nucleotide sequences of the test mRNA might affect its interaction with the ribosome or with other components of the translation apparatus. In particular, consecutive AGG codons generate Shine-Dalgarno sequences, which can interact with the 3′ end of 16S rRNA and which are known to stimulate frameshifting in the RF2 gene (reviewed in reference 2). Inserting five consecutive GGA (Gly) or GAG (Glu) codons introduces nucleotide sequences that are almost identical to five consecutive AGGs, and preliminary results indicate that, although the GGA and GAG codons can both reduce translation somewhat, most of the translational effects of AGG codons reported here must be due to a requirement for translation of the AGG codons.

Of the three sites we used, insertion of AGG codons after codon 13 had the most severe effects on translation. Even two consecutive AGG codons substantially inhibited the synthesis of test protein, and five consecutive codons allowed only a small burst of synthesis (Fig. 2 and 5). A simple explanation is that something needed for translating AGG codons became depleted. The most likely candidate is $tRNA^{AGG}$, which is known to be present in small amounts (9). If the number of test mRNAs being translated became larger than the number of $tRNA^{AGG}$ molecules in the cell, all of the $tRNA^{AGG}$ could potentially become sequestered in peptidyl-$tRNA^{AGG}$-mRNA-ribosome complexes stalled at the first of two consecutive AGG codons (27). If no other process were to release the sequestered $tRNA^{AGG}$, translation of all mRNAs that contain any AGG codons would stop. Such sequestering could not occur if single AGG codons were interspersed with other codons, because translation of the next codon would release the $tRNA^{AGG}$, and indeed, Varenne et al. (26) found no inhibition of translation by interspersed AGG codons.

Although the effects of inserting consecutive AGG codons after codon 13 are entirely consistent with $tRNA^{AGG}$ having become sequestered, the reduction of full-length test protein level was less severe when the same codons were inserted after codon 223 or 307 (Fig. 2 and 5). In addition, novel protein bands began to appear at about the time that test protein synthesis would have stopped had insertion been made after codon 13 (Fig. 2). The mobilities of the novel proteins were consistent with their having been generated by processes such as frameshifting, hopping, or other events leading to chain termination near the AGG codons (2, 10, 19). Such events would release sequestered $tRNA^{AGG}$, which might allow other ribosomes to translate through the consecutive AGG codons to produce full-length protein. Nearly full-length protein might also be generated directly by in-frame hopping to the last AGG codon. The complete loss of ability to synthesize full-length protein when insertion was made after codon 13 argues that the events leading to release of sequestered $tRNA^{AGG}$ or in-frame hopping must occur less frequently at this site than at the other two sites.

These differences in the effects of inserting consecutive AGG codons at codon 13 versus codon 223 or 307 could be explained with one additional hypothesis, namely, that a translating ribosome that has stalled for lack of charged tRNA for the next codon will be stimulated to hop, frameshift, or terminate by the approach of an upstream translating ribosome and thereby release the peptidyl-tRNA. It may be that an upstream ribosome, by translating as close to the stalled ribosome as it can, generates tension on the mRNA between the two ribosomes and increases the probability that the peptidyl-tRNA in the stalled ribosome will disengage from the mRNA, thereby enabling the other events. For this hypothesis to explain our results, either a ribosome stalled at AGG codons 14 to 18 must be close enough to the initiation codon to limit initiation of translation by another ribosome, or stimulation of hopping, frameshifting, or termination must require a nascent peptide of a certain length.

A role for upstream ribosomes in stimulating processes that relieve sequestering could also explain the results of Chen and Inouye (5), who obtained an increased yield of full-length β-galactosidase as five consecutive AGG codons were moved in steps from codon 10 to codon 76. Production of β-galactosidase increased almost eightfold over this interval and reached the level observed when the mRNA contained no AGG codons.

Translation of small amounts of mRNAs produced from the codon test plasmids can be detected in the presence of rifampin, which selectively eliminates host cell mRNAs (Fig. 4). Test and control mRNAs were translated at comparable rates in uninduced cells in the presence of rifampin, suggesting that consecutive AGG codons do not significantly limit translation at sufficiently low levels of mRNA, at least under these conditions.

At the higher levels of mRNA generated after brief induction, translation of AGG-containing test mRNAs increased when measured in the absence of rifampin (Fig. 2) but failed to increase when measured in the presence of rifampin (Fig. 4). The difference might be explained by competition of AGG-containing mRNAs for translation by a limiting amount of $tRNA^{AGG}$. In the absence of rifampin, test mRNA would initially be a small fraction of the total population of AGG-containing mRNAs, and an increase in test mRNA would increase its fractional use of the available $tRNA^{AGG}$ at the expense of the competing mRNAs; in the presence of rifampin, all of the limiting amount of $tRNA^{AGG}$ would be applied to translation of test mRNA, and an increase in the level of this mRNA would have little effect. The undiminished rate of translation of AGG-containing test mRNAs at later times in the presence of rifampin (Fig. 4) argues that the level of translating mRNA did not become high enough to sequester a substantial fraction of the $tRNA^{AGG}$.

A puzzling observation was that the rate of synthesis of control protein failed to increase with mRNA level in only one case, when the test gene contained five consecutive AGG codons after codon 307 and rifampin had been added shortly after induction (Fig. 4). A possible explanation is that ribosomes stalled at the AGG codons might cause translating ribosomes to stack up on the test mRNA. If closely packed, perhaps as many as 30 ribosomes per mRNA could be sequestered ahead of codon 307, half again as many as could be accommodated ahead of codon 223 (20, 25, 28). If essentially all of the functional ribosomes became sequestered in this way, the rate of translation of both test and control mRNAs would be limited by the rate of release of ribosomes from the AGG codons. Ribosome sequestering might also explain the reduced accumulation of control protein observed in the absence of rifampin when five consecutive AGG codons were inserted after codon 223 or 307 of the test gene (Fig. 5).

These interpretations lead to testable predictions. A test for sequestration of $tRNA^{AGG}$ would be to pair the test gene containing five consecutive AGG codons after codon 13 with a control gene containing a single AGG codon. If translation of the test mRNA can sequester essentially all of the $tRNA^{AGG}$, translation of the control mRNA should shut off along with translation of the test mRNA, but the same control mRNA should continue to be translated efficiently when paired with a test gene containing no consecutive AGG codons.

Although it would be difficult to observe products of termination near the string of AGG codons at codons 14 to 18, possible frameshifting could be revealed by introducing a +1 or −1 frameshift immediately downstream of the insertion site; any frameshifting that might occur as $tRNA^{AGG}$ became depleted would then be detected by the appearance of full-length test protein. Moving the AGG codons further from the initiation codon by inserting different numbers of easily translatable codons ahead of them could test the idea that an upstream translating ribosome is needed to stimulate significant levels of frameshifting or hopping.

This codon test system may be useful in determining functional levels of other tRNAs or in probing other limita-

tions of translation capacity in *E. coli*. The ability to sequester ribosomes behind a translational barrier might be exploited for measuring levels of functional ribosomes under different conditions, rates of translation of individual codons, or rates of initiation. Increased understanding of the translation system may also be helpful in optimizing yields of proteins from cloned genes.

ACKNOWLEDGMENTS

We thank Jutta Paparelli for technical assistance, Kathleen Griffin for sequencing plasmids, Eileen Matz for assistance in the Northern analyses, and Paresh Shrimankar for the gift of BL27(DE3).

This work was supported by the Office of Health and Environmental Research of the United States Department of Energy and by grants from the National Institutes of Health (GM27711) and the American Cancer Society (MV-313A).

REFERENCES

1. **Andersson, S. G. E., and C. G. Kurkland.** 1990. Codon preferences in free-living microorganisms. Microbiol. Rev. **54:**198–210.
2. **Atkins, J. F., R. B. Weiss, and R. F. Gesteland.** 1990. Ribosome gymnastics—degree of difficulty 9.5, style 10.0. Cell **62:**413–423.
3. **Bonekamp, F., H. D. Andersen, T. Christensen, and K. F. Jensen.** 1985. Codon-defined ribosomal pausing in *Escherichia coli* detected using the *pyrE* attenuator to probe the coupling between transcription and translation. Nucleic Acids Res. **13:** 4113–4123.
4. **Brinkmann, U., R. E. Mattes, and P. Buckel.** 1989. High-level expression of recombinant genes in *Escherichia coli* is dependent on the availability of the *dnaY* gene product. Gene **85:**109–114.
5. **Chen, G. T., and M. Inouye.** 1990. Suppression of the negative effect of minor arginine codons on gene expression: preferential usage of minor codons within the first 25 codons of the *Escherichia coli* genes. Nucleic Acids Res. **18:**1465–1473.
6. **Dunn, J. J., and F. W. Studier.** 1983. Complete nucleotide sequence of bacteriophage T7 DNA and the locations of T7 genetic elements. J. Mol. Biol. **166:**477–535.
7. **Ikemura, T.** 1981. Correlation between the abundance of *Escherichia coli* transfer RNAs and the occurrence of the respective codons in its protein genes. J. Mol. Biol. **146:**1–21.
8. **Innis, M. A., and D. H. Gelfand.** 1990. Optimization of PCRs, p. 3–12. *In* M. A. Innis, D. H. Gelfand, J. J. Sninsky, and T. J White (ed.), PCR protocols: a guide to methods and applications. Academic Press, Inc., New York.
9. **Komine, Y., T. Adachi, H. Inokuchi, and H. Ozeki.** 1990. Genomic organization and physical mapping of the transfer RNA genes in *Escherichia coli* K12. J. Mol. Biol. **212:**579–598.
10. **Menninger, J. R.** 1983. Computer simulation of ribosome editing. J. Mol. Biol. **171:**383–399.
11. **Panayotatos, N., and K. Truong.** 1985. Cleavage within an RNase III site can control mRNA stability and protein synthesis *in vivo*. Nucleic Acids Res. **13:**2227–2240.
12. **Robinson, M., R. Lilley, J. S. Emtage, G. Yarranton, P. Stephens, A. Millican, M. Eaton, and G. Humphreys.** 1984. Codon usage can affect efficiency of translation of genes in *Escherichia coli*. Nucleic Acids Res. **12:**6663–6671.
13. **Rosenberg, A. H., B. N. Lade, D. Chui, S. Lin, J. J. Dunn, and F. W. Studier.** 1987. Vectors for selective expression of cloned DNAs by T7 RNA polymerase. Gene **56:**125–135.
14. **Rosenberg, A. H., and F. W. Studier.** 1987. T7 RNA polymerase can direct expression of influenza virus cap-binding protein (PB2) in *Escherichia coli*. Gene **59:**191–200.
15. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
16. **Sørensen, M. A., C. G. Kurland, and S. Pedersen.** 1989. Codon usage determines translation rate in *Escherichia coli*. J. Mol. Biol. **207:**365–377.
17. **Sørensen, M. A., and S. Pedersen.** 1991. Absolute *in vivo* translation rates of individual codons in *Escherichia coli*: the two glutamic acid codons GAA and GAG are translated with a threefold difference in rate. J. Mol. Biol. **222:**265–280.
18. **Spanjaard, R. A., K. Chen, J. R. Walker, and J. van Duin.** 1990. Frameshift suppression at tandem AGA and AGG codons by cloned tRNA genes: assigning a codon to *argU* tRNA and T4 $tRNA^{Arg}$. Nucleic Acids Res. **18:**5031–5036.
19. **Spanjaard, R. A., and J. van Duin.** 1988. Translation of the sequence AGG-AGG yields 50% ribosomal frameshift. Proc. Natl. Acad. Sci. USA **85:**7967–7971.
20. **Steitz, J. A.** 1969. Polypeptide chain initiation: nucleotide sequences of the three ribosomal binding sites in bacteriophage R17 RNA. Nature (London) **224:**957–964.
21. **Studier, F. W.** 1991. Use of bacteriophage T7 lysozyme to improve an inducible T7 expression system. J. Mol. Biol. **219:**37–44.
22. **Studier, F. W., and J. V. Maizel, Jr.** 1969. T7-directed protein synthesis. Virology **39:**575–586.
23. **Studier, F. W., and B. A. Moffat.** 1986. Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. J. Mol. Biol. **189:**113–130.
24. **Studier, F. W., A. H. Rosenberg, J. J. Dunn, and J. W. Dubendorff.** 1990. Use of T7 RNA polymerase to direct expression of cloned genes. Methods Enzymol. **185:**60–89.
25. **Takanami, M., and G. Zubay.** 1964. An estimate of the size of the ribosomal site for messenger RNA binding. Proc. Natl. Acad. Sci. USA **51:**834–839.
26. **Varenne, S., D. Baty, H. Verhelj, D. Shire, and C. Lazdunski.** 1989. The maximum rate of gene expression is dependent on the downstream context of unfavourable codons. Biochimie **71:** 1221–1229.
27. **Varenne, S., and C. Lazdunski.** 1986. Effect of distribution of unfavorable codons on the maximum rate of gene expression by an heterologous organism. J. Theor. Biol. **120:**99–110.
28. **Wolin, S. L., and P. Walter.** 1988. Ribosome pausing and stacking during translation of a eukaryotic mRNA. EMBO J. **7:**3559–3569.
29. **Zhang, S., G. Zubay, and E. Goldman.** 1991. Low-usage codons in *Escherichia coli*, yeast, fruit fly and primates. Gene **105:**61–72.